

HUMAN EFFICIENCY FOR VISUAL DETECTION OF TARGETS
ON CRT DISPLAYS USING
A TWO LEVEL MULTIPLE CHANNEL TIME HISTORY FORMAT

By

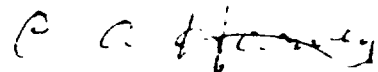
Joseph DiVita, Ph.D

and

Thomas Hanna, Ph.D

NAVAL SUBMARINE MEDICAL RESEARCH LABORATORY
REPORT NUMBER 1101

APPROVED AND RELEASED BY:



C. A. HARVEY, CAPT, MC, USN
COMMANDING OFFICER

NAVAL SUBMARINE MEDICAL RESEARCH LABORATORY

SUMMARY PAGE

PROBLEM

To determine human efficiency for visual detection of targets on CRT displays using a two-level multiple channel time history format.

FINDINGS

Human observers were generally 3-5 dB less sensitive than optimal detectors. Although observers' performance improved as the number of lines in the display increased, this improvement was not as great as that achieved by an optimal detector, particularly for lower signal levels. Less than optimal performance may be due to observers' inability to focus on individual bearings or that they employ suboptimal strategies for using the information at a bearing.

APPLICATION

The finding that observers are 3-5 dB less sensitive than theoretically possible indicates that improved design of visual displays could lead to better detection of sonar targets.

ADMINISTRATIVE INFORMATION

This research was conducted under Naval Medical Research and Development Command Work Unit M0100.001-1022--"Enhanced performance with visual sonar displays." It was submitted for review on 30 March 1987, and designated as NAVSUBMEDRSCHLAB Rep. No. 1101.

PUBLISHED BY THE NAVAL SUBMARINE MEDICAL RESEARCH LABORATORY

ABSTRACT

Human observers were tested for their ability to detect targets on visual displays. The displays simulated the multiple channel time history format of sonar displays, using two levels of intensity encoding. A target was presented on 50% of the trials and appeared as a vertical line at a fixed position. Observers indicated their judgment as to whether a target was present by using a four-category rating scale. Receiver-operating characteristics (ROCs) were generated from the rating data and compared to ROCs for an optimal detector. Data were collected for two signal-to-noise ratios and for 32, 64, and 128 lines of data. Results indicated that observers were 3-5 dB less sensitive than an optimal detector. Performance improved as the number of lines increased but not to the extent predicted by optimal integration of information across lines of the display. Human inefficiency with this display is possibly due to the inability to focus on a single column of data or to the use of a suboptimal decision rule for judging the presence of the target.

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	



INTRODUCTION

Several investigators have experimented with target detectability in the multiple-channel time-history format (MCTH). These displays are used to present data from passive sonar systems and multichannel spectrum analyzers. The Passive Broad Band (PBB) display is an example of this format. In the PBB display, bearing is represented along the horizontal axis, time along the vertical axis, and amplitude of the signal is encoded by pixel intensity. Currently, these displays are monochromatic, and signal strength may be mapped onto two, four, or eight luminance levels (also referred to as brightness levels).

Thompson (1975) has demonstrated that target detectability improves as the number of luminance levels used to display information increases. There is also evidence that displaying information in color as opposed to monochromatic luminance levels may improve target detectability (Evans 1968). Although these experiments have provided empirical evidence that operator performance varies with display format, no comparison between actual human performance and ideal performance has been made, nor have the investigators attempted to specify those critical features in the displays that lead to improved performance.

Delisle and Kroenert (1983) have proposed a machine based, analytical model of performance detection. Given certain constraints placed upon the target, this model represents an ideal detector. The purpose of our research is to evaluate operator performance in terms of ideal performance.

Delisle and Kroenert's model assumes that the display is composed of two brightness levels, on and off. The probability that noise-only data will appear on the screen is 0.5. Thus noise-only data will appear as a "random dot" display. The

probability that the target data will appear on the screen is greater than 0.5. How much greater than 0.5 depends on the target's strength; the stronger the target's signal the more likely its data will appear on the screen. If we assume that a target's bearing is constant, that is, the target's horizontal position in the display remains constant over a specified number of temporal updates of data, then the ideal detector adjudicates whether a target is present, based on the number of lit cells (or bins) in a specified column. The ideal detector sets a criterion for detection: If j or more cells are lit over a specified number of lines, then a detection is made. If these cells were lit by noise, the judgment is called a false alarm; if they were lit by a target, the judgment is called a hit. When L lines of data are presented, the probability that j or more cells will be lit by either the noise or target can be computed from the binomial distribution:

$$(1) \quad P_{(j)} = \sum_{i=j}^L L! / (i! * (L-i)!) * p^i * q^{L-i}$$

Where

$P_{(j)}$: is the probability that at least j cells are lit.

p : is the probability of a cell being lit.

q : is the probability of a cell being off.

L : is the number of lines in the display.

i : index referring to detection criterion.

As mentioned above, for the noise distribution, $p = q = 0.5$.; whereas for the signal distribution, $p > 0.5$, $q = 1-p$.

By plotting the respective probabilities from Eq. 1 for noise versus target as j varies from 0 to L , we trace out the ROC curve

(Green and Swets 1966) for the two distributions. Thus for a given false alarm rate, the hit rate can be calculated, and these two rates correspond to the performance of the ideal detector. The performance of this detector is ideal in the sense that for a given false alarm rate, it is impossible to obtain a higher hit rate. Figure 1 depicts the ROC curves for a signal strength of -2 db and -5 db across 32, 64 and 128 lines of data. The probabilities have been converted to z-scores, yielding lines whose slopes are all approximately equal to 1.

The curves in Figure 1 represent ideal performance. That is, for a given number of lines of data, associated with each false alarm rate (the x coordinate) is the best possible hit rate, (the y coordinate). If human performance is equivalent to ideal performance, then the observer's false alarm rate and hit rate, taken as a (x,y) paired coordinate, should plot onto the respective ROC curve. In this manner, an observer's performance may be compared to ideal performance. Furthermore, by having the subject give a confidence judgment, performance can be compared to ideal at a variety of false alarm rates - in essence, tracing out the ROC for the human observer by having them vary their criteria for detection.

One consequence of Delisle and Kroenert's model is that performance should improve as the number of lines in the display increases (see Figure 1). That is, the more data available to the ideal detector, the more accurate its judgment will be. The current research evaluates human performance for displays of 32, 64, and 128 lines of data, in order to determine the degree to which the human observer integrates information when presented with more lines of information.

Delisle and Kroenert's model assumes that the observer is only asked to judge whether or not a target is present at one

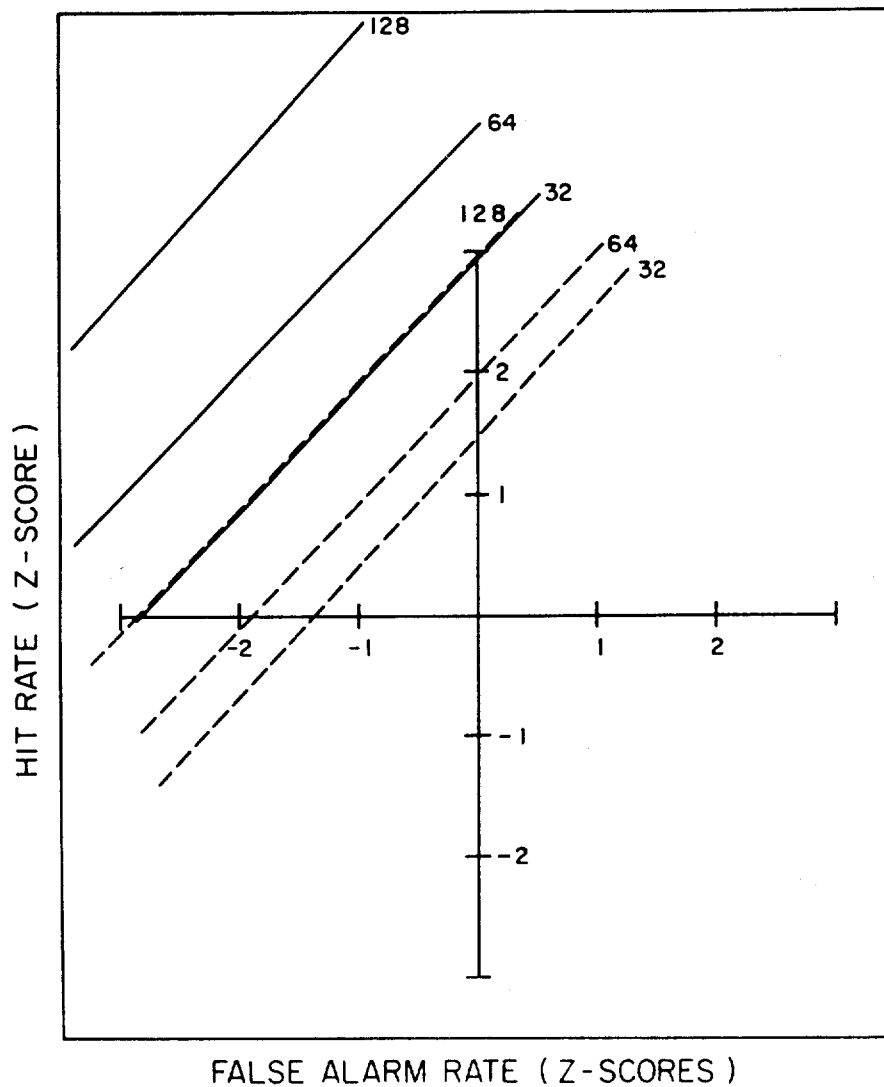


Figure 1. ROC curves for ideal performance. The probabilities have been converted to z-scores. Solid line: SNR of -2 db. Dashed line: SNR of -5 db.

bearing. Thus in their model there is no uncertainty as to the position of the target, only whether or not a target is present. In this sense the model is strictly a detection model. In practice however, with such displays, the observer typically has to judge not only whether or not a target is present but also must decide where the target is located, i.e. at what bearing the target is coming in. Thus, the task generally entails both detection and identification (identification in this context refers to correctly specifying the bearing). In our experiments, the subjects had only to detect the presence of the target; its location was specified.

METHOD

Subjects

Three members of the Vision Department at The Naval Submarine Medical Research Laboratory were subjects. Two were experienced with waterfall displays.

Apparatus

A Vax 750, a Ramtek 9400 graphics display generator, and Matsushita standard phosphor color monitor were used to simulate the sonar display. The addressability of the monitor was 1024 by 1280 pixels (100 pixels per inch). The C.I.E. chromaticity coordinates (x,y) of the phosphors were 0.60, 0.34 for the red, 0.28, 0.59 for the green, and 0.16, 0.07 for the blue. The display was illuminated by two fluorescent tubes located behind and above the observer. The lamps were covered with neutral density filters which reduced the illumination falling on the screen to 0.25 fc, as measured by a Gossen light meter.

Display

The display simulated one depression-elevation sector of a spherical array passive broadband (SAPBB) short term averaging (STA) display, with bearing represented along the horizontal axis, time along the vertical axis, and amplitude of signal encoded by pixel intensity. The display may be conceptualized as a rectangular matrix. The number of columns, i.e. bearings, was always 60, whereas the number of rows, i.e. temporal updates, was an experimental variable. Displays of 32, 64, and 128 lines of data were tested. At a viewing distance of 2 ft, these three displays subtended visual angles of 2.34×4.77 deg, 4.47×4.77 deg, 8.93×4.77 deg, respectively. A white cross hair was placed directly above the 31st bearing, at approximately the middle of the display.

One bin of information, which corresponds to one element in the array, was represented by a 3×3 block of pixels. Only two luminance levels, "on" and "off", were used to encode pixel intensity. The on state utilized the green phosphor, and corresponded to a luminance value of .35 fL.

The random energy in the sea may be simulated by a Gaussian distribution with a mean of zero and a standard deviation of one. The mapping of noise to luminance level, in this case "on" or "off", entails the arbitrary assignment of values sampled from the noise population to luminance level. In this display, all values less than or equal to zero mapped onto a luminance level of screen off, whereas any value greater than zero mapped onto screen on. That is, the average marking density was 0.5. The signal distribution is a Gaussian with variance equal to that of the noise (i.e. S.D.=1). However, the mean of the signal's distribution, M_s , is shifted from that of the noise, M_n , depending on the strength of the signal. The mean of the signal

distribution is given by the equation:

$$2) \quad M_s = M_n + 10 \text{ SNR}/10$$

Signal strengths (SNRs) of -2 db, and -5 db were tested in the experiment. These values of SNR correspond to average marking densities of 74% and 63%, respectively.

Procedure

In the experiment, one trial consisted of the following events. The display waterfalled at a rate of approximately 10 lines/sec until an entire new block of data was presented, and then stopped. The subject was instructed to judge whether a target was present at the 31st bearing only. (This bearing was marked by a cross hair.) Subjects rated the confidence of their responses with the following scale: 1 - indicated that they were relatively certain that a target was not present, 2 - no target but some uncertainty, 3 - target present but some uncertainty, 4 - relatively certain that a target was present. Subjects had an unlimited amount of time in which to view the static display and make their decision. After entering their response, a new trial started.

The probability that a target was presented on any one trial was 50%. As mentioned, the subjects were tested on displays of 32, 64, and 128 lines of data. The number of lines of data in the display remained constant for 100 trials. Each subject completed 5 blocks of 100 trials at each of the two signal levels and for each of the three number of lines tested. Across blocks of trials the order of presentation of both the number of lines and signal strength was randomized. Subjects knew whether they were searching for a weak or strong target in each block of trials.

RESULTS

The four response categories were grouped to yield three points on the ROC curve in the following manner. First, a lax criterion was generated by considering a response of 1 as no target present, and a response of 2, 3, or 4 as target present. Second, a neutral criterion was generated by grouping response 1 and 2 as no target present, and response 3 and 4 as target present. Lastly, a strict criterion was obtained by grouping responses 1, 2, and 3 as a no response and 4 as a yes response. The results for the 32, 64 and 128 line displays are plotted separately for each subject in Figures 2, 3, and 4. Unfortunately, we were unable to obtain a data point generated by the strict criterion across all three number of lines tested for the -2 db target for one subject. This subject had no false alarms that were rated a 4 at this signal strength.

As is illustrated in Figures 2, 3, and 4, we were able to obtain coherent ROC curves for each subject. Ideal performance for a signal strength of -2, -5, and -8 db has been plotted for reference. To further analyze the results, best fitting lines were fit to the data points. Tables 1, 2, and 3 list the slopes, average slopes and intercepts of these lines. The slopes of these lines are all approximately equal to one as the model predicts. In fact, the data are rather consistent. In general, the subjects performed approximately 3 db below ideal performance. (Subject LB, who was previously unfamiliar with waterfall displays, is almost 5 db below ideal performance.) That is, subject performance on the -2 db target was what one would expect for a weaker target of -5 db, whereas performance on the -5 db target had degraded to performance expected for a signal strength of -8 db. Tables 4 and 5 demonstrate this point. For given false alarm rates of 10% (Table 4) and 5% (Table 5),

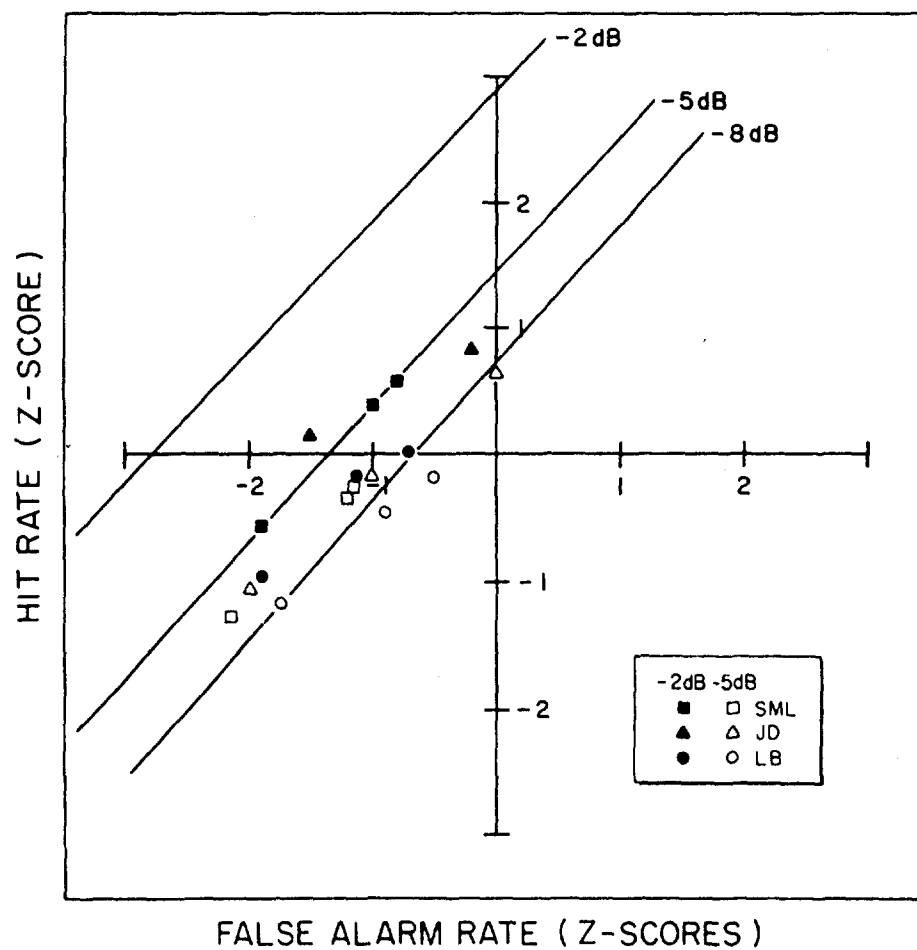
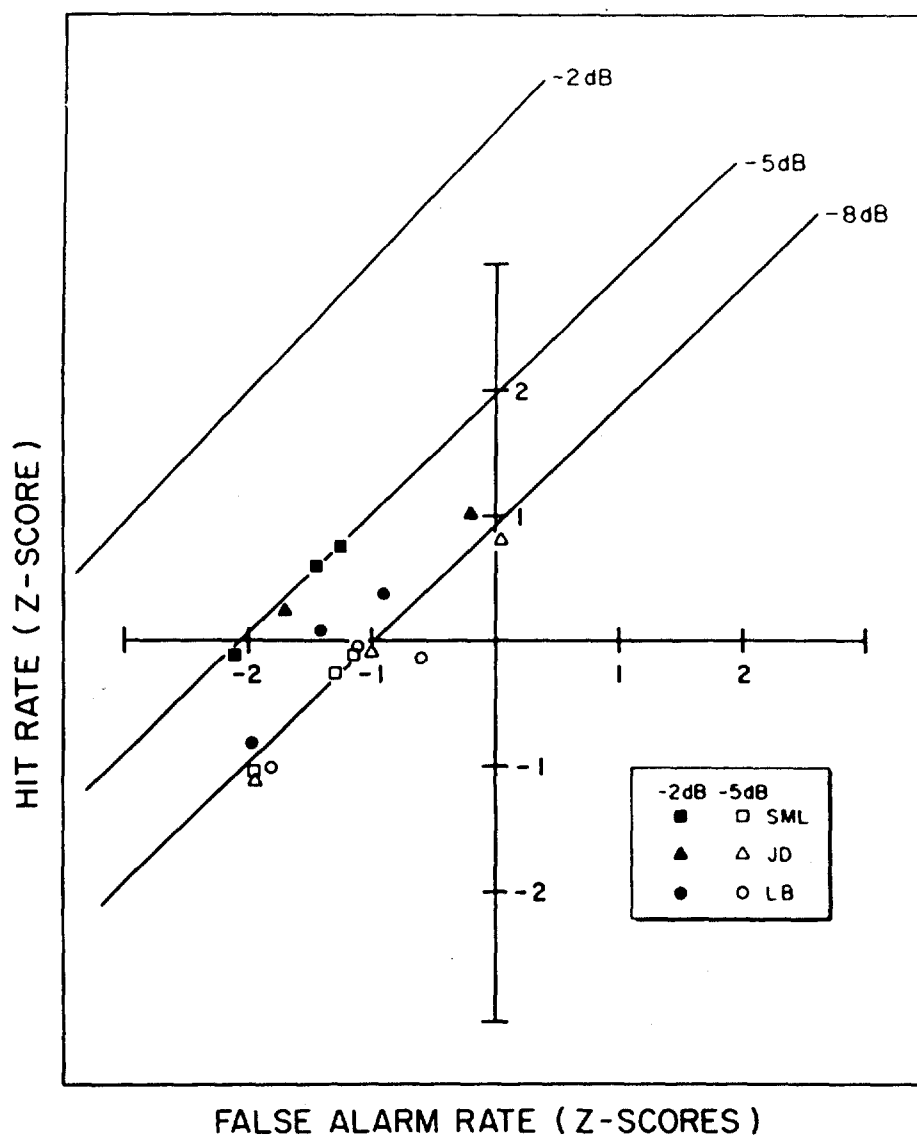


Figure 2. ROC curves for each subject when tested on 32 lines of data. The ROC curve for Ideal performance is given for reference by the solid lines for signal strengths of -2, -5, and -8 db.



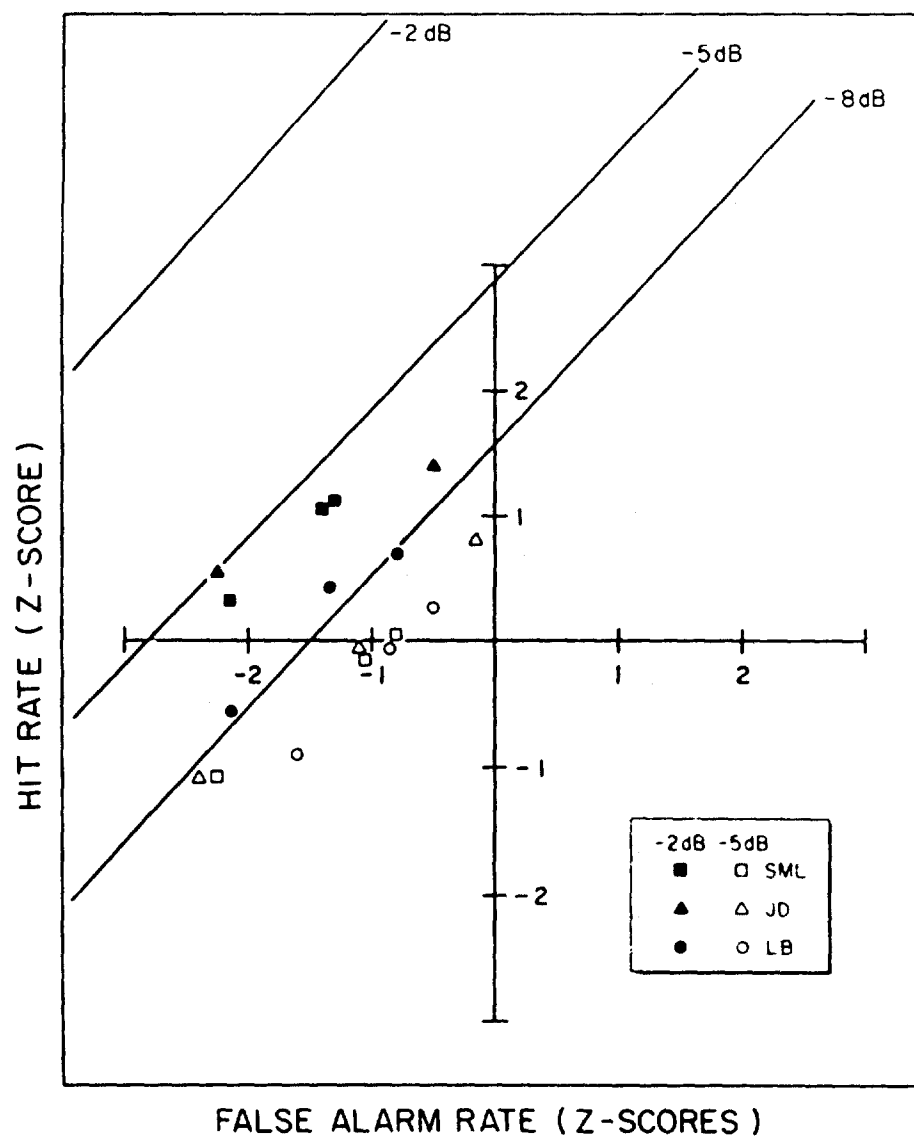


Figure 4. ROC curves for each subject when tested on 128 lines of data. The ROC curve for Ideal performance is given for reference by the solid lines for signal strengths of -2, -5, -8 db.

TABLE 1: Slopes of best fitting line for each subject's ROC.

		SML	JD	LB
32	-2	.993	.507*	.884
	-5	.998	.87*	.796
64	-2	1.044	.524*	1.188
	-5	1.102	.939	.854
128	-2	.978	.469*	1.002
	-5	.792	.837	1.08

* Only two points on the ROC were obtained.

TABLE 2: Mean slopes of the ROC curves.*

	-2 db	-5 db
32	.939	.889
64	1.116	.965
128	.990	.903

* Lines generated by only two points have been excluded.

TABLE 3: Y intercepts in terms of z scores for best fitting lines for each subject's ROC in each condition.

		SML	JD	LB
32	-2	1.44	1.63	0.76
	-5	0.88	0.71	0.24
64	-2	2.17	1.18	1.59
	-5	1.16	0.76	0.57
128	-2	2.43	0.95	1.63
	-5	0.71	0.92	0.84

TABLE 4: Hit rates, in percent, from a 10% false alarm rate ($z = -1.28$) extrapolated from best fitting lines for each subject in each condition.

		IDEAL	SML	JD	LB
32	-2	94.14	56.55	61.94	35.61
	-5	56.32	34.61	34.20	21.77
	-8	28.77			
64	-2	99.69	78.23	69.64	52.75
	-5	77.40	40.25	33.07	29.87
	-8	39.63			
128	-2	99.99	88.0	84.92	63.57
	-5	94.66	38.13	43.84	29.32
	-8	56.55			

TABLE 5: Hit rates, in percent, from a 5% false alarm rate ($z = -1.65$) extrapolated from best fitting lines for each subject in each condition.

		IDEAL	SML	JD	LB
32	-2	88.06	42.0	54.66	24.35
	-5	41.25	22.21	23.27	14.14
	-8	17.52			
64	-2	99.08	65.32	62.55	35.53
	-5	64.65	25.62	21.59	19.93
	-8	26.24			
128	-2	99.99	79.19	80.48	49.08
	-5	89.07	27.59	32.10	17.28
	-8	41.68			

each subject's hit rate, extrapolated from the best fitting line, may be compared to the ideal hit rate. For example, with 32 lines, a signal level of -2 db, and a false alarm rate of 10% (Table 4), obtained hit rates range from 36% to 62%. This percentage is much less than the 94% possible for the ideal detector with a signal level of -2 db, but corresponds well to the 56% for the ideal detector with a signal level of -5 db.

Table 6 shows the equivalent loss in SNR for each condition relative to ideal performance. The SNR losses were obtained in the following manner. Values of d'^* (Green and Swets, 1966/1974) were estimated at a false alarm rate of 5% for the ideal detector and the three subjects. Since, for the ideal detector, d' is approximately proportional to signal energy, SNR loss can be computed as:

$$3) \quad 10 \log(d'_{\text{obs}}/d'_{\text{ideal}})$$

where d'_{obs} is the d' for the observer and d'_{ideal} is the d' for the ideal observer. The values for the SNR loss range from -1.9 to -6.1 db. Table 7 gives the same information as Table 6 except averaged across subjects. Tables 6 and 7 show that as the number of lines increased, the discrepancy between subject performance and ideal performance increases. In fact, for the 128 line display the discrepancy is clearly greater than -3 db for all subjects.

* d' is a measure of sensitivity defined as the difference between the mean of two Gaussian distributions divided by their common standard deviation.

Table 6: d' and SNR loss for each subject in each condition for a false alarm rate of 5%.

		d'				SNR LOSS		
		IDEAL	SML	JD	LB	SML	JD	LB
32	-2	2.83	1.45	1.77	.95	-2.9	-2.0	-4.7
	-5	1.43	.88	.92	.58	-2.1	-1.9	-3.9
64	-2	4.01	2.04	1.97	1.28	-2.9	-3.1	-5.0
	-5	2.03	.99	.86	.81	-3.1	-3.7	-4.0
128	-2	5.68	2.46	2.51	1.63	-3.6	-3.5	-5.4
	-5	2.88	1.05	1.18	.71	-4.4	-3.9	-6.1

TABLE 7: Average d' and SNR loss at a false alarm rate of 5% for the three subjects. Ideal d' is given for reference.

	d'			SNR LOSS	
	-5	-2		-5	-2
32	.79	1.39	32	-2.6	-3.2
64	.89	1.76	64	-3.6	-3.7
128	.98	2.20	128	-4.8	-4.2

	IDEAL d'	
	-5	-2
32	1.43	2.83
64	2.03	4.01
128	2.88	5.68

Although human performance was relatively poor as compared to ideal performance, in general as the number of lines increased in the display, subjects' performance improved. This result is in accordance with Delisle and Kroenert's model and is clearly illustrated in Figure 5 where SML's data are plotted for the 32, 64 and 128 line display. However, as noted above, this improvement is not as large as expected, particularly for an SNR of -5 db where only a small improvement is observed as a function of the number of lines (see Figure 6). Table 7 also shows the small improvement obtained at an SNR of -5 db: average d' only increases from .79 to .98 as the number of lines increases from 32 to 128.

Although subjects were given an unlimited amount of time to view the static display and make their decision, all subjects noted that their decision had been made by the time the display had stopped waterfalling. Thus the decision as to whether a target was present was based on a spontaneous perception of the display as opposed to careful scrutiny and counting of individually lit bins of data.

DISCUSSION

These results show that human performance is very poor when compared to ideal performance. For example, if we assume cylindrical spreading and minimum loss due to absorption, a 3 db loss means that an ideal detector would pick up a contact at twice the distance that the human detector is capable of. If we are at all concerned with the sonar operator's ability to detect targets, we must determine why performance is so poor. We may then systematically vary the manner in which information is presented and test whether these changes improve operator performance.

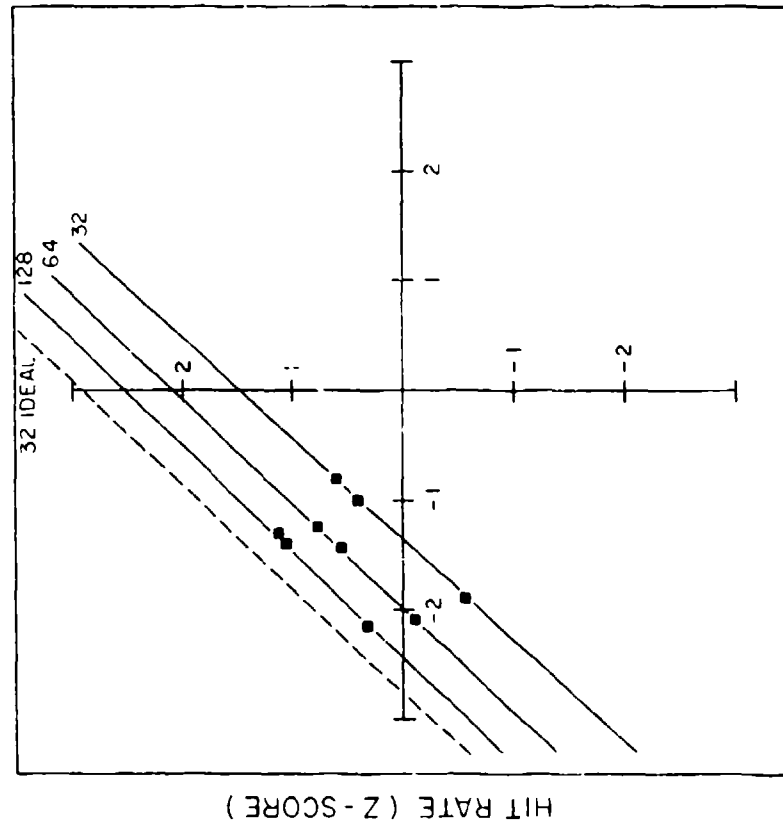


Figure 5. ROC curves, solid lines, for subject SML on the -2 db target for 32, 64 and 128 lines of data. As Delisle and Kroenert's model predicts, performance improves as the number of lines increases; however, performance is far below the ROC curve for ideal performance with only 32 lines of data (broken line).

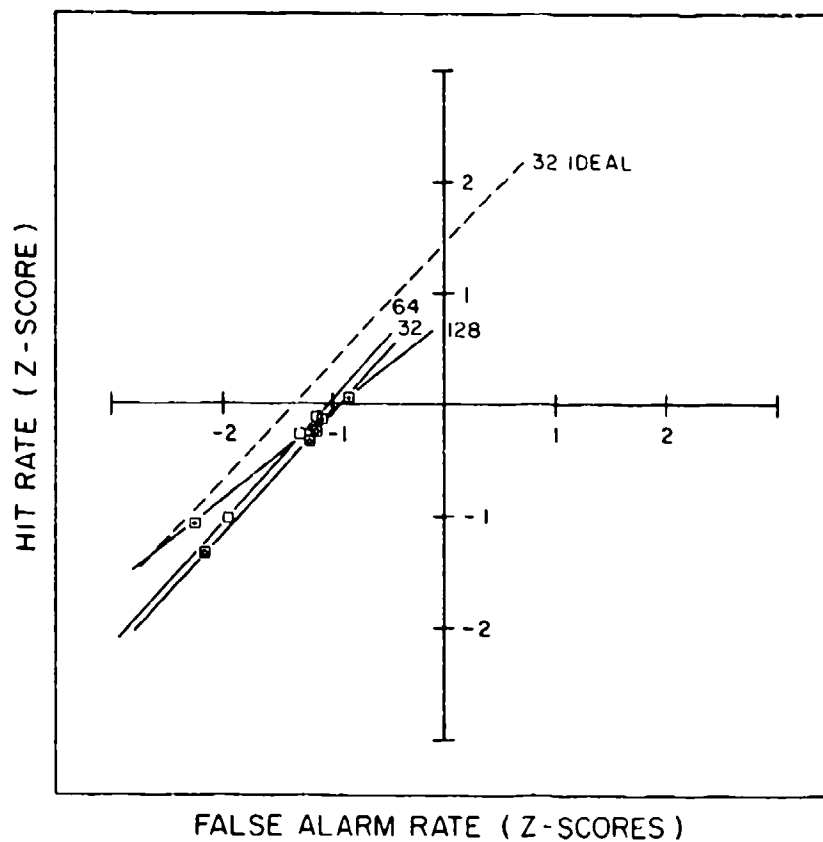


Figure 6. ROC curves, solid lines, for subject SML on the -5 db target for 32, 64 and 128 lines of data. Performance improves only slightly as the number of lines increase from 32 to 64; with 128 lines, the slope of the ROC is less than one, making it difficult to ascertain if there has been further improvement. Note that performance is far below the ROC curve for ideal performance at only 32 lines of data (broken line).

One factor that may be detrimental to performance is uncertainty (Green and Swets 1966/1974). Despite the fact that the target's bearing was clearly marked by a cross hair, the subjects still experienced some uncertainty as to which column was the target column. This is because the random dot pattern of the noise makes it difficult, if not impossible, to clearly discern a single column of the display from top to bottom. Thus, the subject may be looking across several columns in the immediate vicinity of the target column. This would have a detrimental effect on performance, because the probability that one of several noise-only bearings may appear to contain a signal is greater than the probability that the single target bearing will appear to contain a signal when noise is presented at that bearing. Thus the subject is more likely to confuse noise for signal. The extent to which uncertainty is degrading performance must be evaluated directly in a separate experiment.

Another factor which may be detrimental to performance is the reliance on the "clustering" of lit pixels in order to detect targets. It is apparent that a string of lit pixels clustered together stands out from the background as a short line and appears to be a target. It is generally true that targets will exhibit more clustering than noise; however, given a fixed marking density, i.e. a given number of lit pixels over a given number of lines, the probability that these data are from a signal of a particular strength is independent of the clustering of the lit pixels within that string. Thus, if we have two data strings with equal marking density, as illustrated in Figure 7, subjects should, ideally, treat these strings as equally likely to contain a target. However, they may be more apt to claim that string A is a target because the clustering of lit pixels is greater in string A than B. They are thus adopting a suboptimal strategy and their performance departs from ideal performance.



Figure 7: Illustration of the grouping factor. The marking densities of strings A and B are equal; however, the clustering factor is greater in string A than in string B.

The extent to which this clustering factor has an adverse effect on target detection must also be analyzed by obtaining ROC curves which assume different degrees of clustering.

REFERENCES

1. Thompson, W.G., Passive sonar display studies - FY'75. NUC TP 487, San Diego: Naval Under Sea Center, Sept 1975 (DTIC AD no B008231L)
2. Evans, T.R., The Assement of Color as a Sonar Display Parameter. Doc No. 68-766-U, Austin TX: Tracor Sciences and Systems, June 1968.
3. Butler, W.B., McKemie, W.M., Engineering Guidelines for the use of Color in Sonar a Display. Doc No. T73-AU-9550-U (Rev), Austin TX: Tracor Sciences and Systems, April 1974.
4. Delisle, R.B. and Kroenert, J.T., An Analytical Model for the Detection Performance of Multiple Channel Time History Display Formats. Journal of the Acoustical Society America, 73 (6), June 1983.
5. Green, D.M., Swets, J.A., Signal Detection and Psychophysics. Robert E. Krieger Publishing Co., Huntington NY, 1974.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NSMRL Rpt 1101	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Human Efficiency for Visual Detection of Targets on CPT Displays Using a Two Level Multiple Channel Time History Format		5. TYPE OF REPORT & PERIOD COVERED Interim report
7. AUTHOR(s) Joseph DiVITA and Thomas HANNA		6. PERFORMING ORG. REPORT NUMBER NSMRL Rpt 1101
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Submarine Medical Research Laboratory Naval Submarine Base New London Groton, CT 06349-5900		8. CONTRACT OR GRANT NUMBER(s)
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Medical Research & Development Command NMCNCR, Bethesda, MD 20814-5044		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS M0100-001.1022
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE 2 Oct 1987
		13. NUMBER OF PAGES 21
		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) ROC curves; ideal detector; MCTH; pbb		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) A Human observers were tested for their ability to detect targets on visual displays. The displays simulated the multiple channel time history format of sonar displays, using two levels of intensity encoding. A target was presented on 50% of the trials and appeared as a vertical line at a fixed position. Observers indicated their judgment as to whether a target was present by using a four-category rating scale. Receiver-operating characteristics (ROCs) were generated from the rating data and compared to		

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ROCs for an optimal detector. Data were collected for two signal-to-noise ratios and for 32, 64, and 128 lines of data. Results indicated that observers were 3-5 dB less sensitive than an optimal detector. Performance improved as the number of lines increased, but not to the extent predicted by optimal integration of information across lines of the display. Human inefficiency with this display is possibly due to the inability to focus on a single column of data or to the use of a suboptimal decision rule for judging the presence of the target.

Best Available Copy

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)